

ARCHXAI – AI ENHANCED CROSSBORDER ARCHIVES PRIVACY STATEMENT

Data Protection Act 2018/1050, EU General Data Protection Regulation 2016/679.

This document clarifies and specifies the information on the processing of personal data on a research or project-by-project basis. The privacy statement helps ensure the data controller's obligation to demonstrate compliance. Demonstrating compliance is a key principle in the data protection regulation, meaning that the data controller must be able to prove that they adhere to data protection act.

1. Registrar

South-Eastern Finland University of Applied Sciences Oy / Kaakkois-Suomen
ammattikorkeakoulu Oy
PL 68 (Patteristonkatu 3), 50101 Mikkeli, Finland
Registration number: 2472908–2
<https://www.xamk.fi/>

Joint registrars:

The National Archives of Finland / Kansallisarkisto
Rauhankatu 17, 00171 Helsinki, Finland
Y-tunnus: 0245422-3
<https://kansallisarkisto.fi/>

National Archives of Estonia / Rahvusarhiiv
Nooruse 3, 50411 Tartu, Estonia
no registration number
<https://www.ra.ee/>

The National Archives of Latvia / Latvijas Nacionālais arhīvs
Šķūņu street11, Riga, LV-1050, Latvia
Registration number: 90009476367
<https://www.arhivi.gov.lv/lv>

2. Contact person for register matters

Anna Ollanketo, projektipäällikkö
Kaakkois-Suomen ammattikorkeakoulu (Xamk)
Patteristonkatu 2, 50100 Mikkeli
+358 401839330
anna.ollanketo@xamk.fi

Ilkka Jokipii
National Archives of Finland

ilkka.jokipii@kansallisarkisto.fi
+358 295337341

Tarvo Kärberg
National Archives of Estonia
tarvo.karberg@ra.ee
+372 51906687

Mairita Lukianska
National Archives of Latvia
mairita.lukianska@arhivi.gov.lv
+371 2726 3189

3. What is the purpose for which my personal data is processed?

Scientific research is one of Xamk's core tasks. Each research project will separately identify the need for the processing of personal data and will inform the subjects separately.

Personal data is collected in surveys and events, such as workshops, carried out within the project in order to contact participants for providing or obtaining additional information and to support the development of cooperation.

The processing of personal data is based on the information provided by the participant when taking part in the project activities, as well as on the data collected for the funder in accordance with the funder's guidelines.

4. On what basis is my personal information processed?

The processing criteria will be defined separately for each research project and will be communicated to the subjects.

The processing of personal data in scientific research is based on the general EU data protection regulation as well as national data protection law. Scientific research is a task in the public interest within the meaning of the Data Protection Regulation, in which case scientific research in the public interest is the basis for the processing of personal data (Data Protection Act). In the Data Protection Regulation (recital 159), the concept of research also covers RDI activities.

5. Name, nature, duration and performers of the study or project

Name of study: ArchXAI - AI enhanced crossborder archives
Nature of the study: development project
Duration of the study: 1.6.2025-31.5.2028

The persons conducting the research: Project personnel at Xamk and the participating National Archives

The ArchXAI project is funded by the Interreg Central Baltic programme.

The project is carried out as a joint initiative between Xamk and the National Archives of Finland, Estonia and Latvia from 1 June 2025 to 31 May 2028.

6. Parties to the collaborative study and division of responsibilities (delete if not a collaborative project)

Xamk: Acts as the lead partner, responsible for overall project coordination, management, and reporting. Xamk develops the AI-assisted toolset for archivists and end-users, facilitates project communication, and together with the partners organises project events such as workshops, stakeholder and feedback surveys, and dissemination of results.

The National Archives of Finland (NAF): Responsible for the digitisation of archival collections, including pre- and post-processing, as well as developing two image-to-text tools for archival use. The partner conducts stakeholder and feedback surveys, manages project communication and dissemination, and together with the consortium, coordinates and reports on project activities, organises events such as workshops, and disseminates project results.

National Archives of Estonia (NAE): Designs and implements algorithms, NLP methods, and fine-tunes AI models for analysing born-digital materials and produces an open-source AI-based indexing and cataloguing tool. The partner is responsible for digitising archival collections, conducting stakeholder and feedback surveys, managing project communication and dissemination, coordinating and reporting on activities, organising events such as workshops, and disseminating project results.

The National Archives of Latvia (NAL): Responsible for the digitisation of archival collections and for adopting and implementing the applications developed within the project. The partner is responsible for digitising archival collections, conducting stakeholder and feedback surveys, managing project communication and dissemination, coordinating and reporting on activities, organising events such as workshops, and disseminating project results.

7. What kind of information about me is processed?

For maintaining stakeholder and customer relations in the project, the necessary personal data include the person's name, organisation, and contact details (phone number, email address). The contact information is used for project-related communication, marketing of workshops and events, and conducting stakeholder and feedback surveys. Personal data are collected only to the extent necessary for implementing the project.



Co-funded by
the European Union

Central Baltic Programme



The research does not involve processing special categories of personal data as defined in Article 9 of the GDPR (sensitive personal data).

The RDI work carried out in the project relates to the processing of large digital datasets, the use of artificial intelligence, and the development of operational practices within the participating organisations through these solutions. The project includes, among other things, testing the further training of language models underlying AI systems. In cases where project partners provide materials that may contain personal data for the purposes of method development, the handling and protection of such data will, where needed, be agreed upon in a separate contract between the parties, specifying the conditions for personal data processing.

8. From what sources is my information collected?

The data are collected directly from the data subjects with their explicit consent. The information is gathered through forms (surveys, registration and participation forms for events and workshops), interviews and by email. In addition, data are collected through interviews and other research methods, such as observation, to support application development.

9. Will my personal information be disclosed to third parties?

The data are not regularly disclosed to third parties. Information may be published only to the extent agreed with the data subject. Certain data may be disclosed to the project's funder as part of mandatory reporting requirements, such as participant lists for events.

10. Is my data processed outside the EU or the EEA?

The data are not transferred by the controller outside the EU or the EEA. The data processed in the project is handled on on-premise hardware and within secure HPC platforms inside EU. Only people who have been specifically granted access may use the data.

At Xamk, Microsoft cloud services (Teams and OneDrive) are used as collaboration platform.

Microsoft is committed to complying with the EU General Data Protection Regulation. Microsoft's Privacy Statement is available at: <https://privacy.microsoft.com/fi-FI/privacystatement>

11. How long will my data be kept?

Personal data is retained for the duration of the project and for as long as they are needed for their intended purpose. In project activities, personal data are stored according to the archiving period required by the funders. Project materials and datasets, which may also include personal data, are archived in accordance with Xamk's practices: for EU-funded projects, the retention obligation is 10 years after the end of the programme period.

12. How is my personal information protected?

The data of individuals participating in the project are stored in information systems that require personal user credentials. Only those personnel of the data controller who have a work-related right to access the data can view and use the information in the system. Project staff are bound by a duty of confidentiality and are required to keep confidential any information obtained in connection with the processing of personal data.

Customer and stakeholder information is stored in a customer management system, where data maintenance is protected by an AD account and password. The information stored in the system is accessible only to a restricted group of users with defined access rights.

The security and confidentiality of the personal data register are ensured through appropriate technical and administrative measures, in accordance with good data processing practices.

13. How can I exercise my rights under the Privacy Regulation?

The rights of the register are listed in more detail in the research information data group's privacy statement (xamk.fi/tietosuojailmoitus).

Subjects / registrants can make all requests related to this study to the contact person below.

Anna Ollanketo, Project Manager
South-Eastern Finland University of Applied Sciences (Xamk)
Patteristonkatu 2, 50100 Mikkeli
+358 401839330
anna.ollanketo@xamk.fi

When the legal basis for processing is public interest, the data subject has the following rights:

- the right to receive information about the processing of personal data, unless an exception is expressly provided by law
- the right to access their data
- the right to rectify data
- the right to restrict the processing of data
- the controller's obligation to inform about rectification or restriction of processing
- the right to object to the processing of data
- the right not to be subject to automated decision-making without a lawful basis